SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Ryals, John
                Delaney, Terry
                Friedrich, Leslie
                Weymann, Kristianna
                Lawton, Kay
                Ellis, Daniel
                Uknes, Scott
                Jesse, Taco
                Vos, Pieter

    (ii) TITLE OF INVENTION: GENE ENCODING A PROTEIN INVOLVED IN THE SIGNAL TRANSDUCTION CASCADE LEADING TO SYSTEMIC ACQUIRED RESISTANCE IN PLANTS

    (iii) NUMBER OF SEQUENCES: 17

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Novartis Corporation
        (B) STREET: 520 White Plains Road, P.O. Box 2005
        (C) CITY: Tarrytown
        (D) STATE: New York
        (E) COUNTRY: USA
        (F) ZIP: 10591

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Meigs, J. Timothy
        (B) REGISTRATION NUMBER: 38,241
        (C) REFERENCE/DOCKET NUMBER: CGC 1909

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (919) 541-8587
        (B) TELEFAX: (919) 541-8689

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9919 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

76

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
TGATCATGAA TTGCGTGTAG GGTTGTGTTT TAAAGATAGG GATGAGCTGA AGAAGGCGGT      60

GGACTGGTGT TCCATTAGAG GGCAGCAAAA GTGTGTAGTA CAAGAGATTG AGAAGGACGA     120

GTATACGTTT AAATGCATCA GATGGAAATG CAATTGGTCG CGTCGGGCAG ATTGAATAGA     180

AGAACATGGA CTTGTTAAGA TAACTAAGTG TAGTTGGTCC ACATACTTGT TGTTCTATTA     240

AGCCGGAAAA CTTCAACTTG TAATTTGCAG CAGAAGAGAT TGAGTGTCTG ATCAGGGTAC     300

AACCCACTCT AACAGCAGAG TTGAAAAGTT TGGTGACATG CTTAAAACTT CAAAGCTGCG     360

GGCAGCAGAA CAGGAAGTAA TCAAAGATCA GAGTTTCAGA GTATTGCCTA AACTAATTGG     420

CTGCATTTCA CTCATCTAAT GGGCTACTTG TGGACTGCAA TATGAGCTTT TCCCTAATCC     480

TGAATTTGCA TCCTTCGGTG GCGCGTTTTG GGCGTTTCCA CAGTCCATTG AAGGGTTTCA     540

ACACTGTAGA CCTCTGATCA TAGTGGATTC AAAAGACTTG AACGGCAAGT ACCCTATGAA     600

ATTGATGATT TCCTCAGGAC TCGACGCTGA TGATTGCTTT TTCCCGCTTG CCTTTCCGCT     660

TACCAAAGAA GTGTCCACTG ATAGTTGGCG TTGGTTTCTC ACTAATATCA GAGAGAAGGT     720

AACACAAAGG AAAGACGTTT GCCTCGTCTC CAGTCCTCAC CCGGACATAG TTGCTGTTAT     780

TAACGAACCC GGATCACTGT GGCAAGAACC TTGGGTCTAT CACAGGTTCT GTCTGGATTG     840

TTTTTGCTTA CAATTCCATG ATATTTTTGG AGACTACAAC CTGGTGAGCC TTGTGAAGCA     900

GGCTGGATCC ACAAGTCAGA AGGAAGAATT TGATTCCTAC ATAAAGGACA TCAAAAAGAA     960

GGACTCAGAA GCTCGGAAAT GGTTAGCCCA ATTCCCTCAA AATCAGTGGG CTCTGGCTCA    1020

TGACCAGTGG TCGGAGATAT GGAGTCATGA CGATAGAAAC AGAAGATTTG AGGGCAATTT    1080

GTGAAAGCTT TCAGTCTCTT GGTCTATCAG TGACAGCGAA CGCACCTGCA CATGTGGGAA    1140

GTTTCAATCG AAGAAGTTTC CATGTATGCA CCCAGAAATG GTGCAAAGGA TTGTTAACTT    1200

GTGTCATTCA CAAATGTTGG ATGCAATGGA GCTGACTAGG AGAATGCACC TTACACGCCC    1260

ACTCAGTGTT CTCTTATCTC TAGACCTGAA ACTAACTTGC TGTGTAATTC GAGTTACAAA    1320

AGGTTAAAGG AAGAATTAGG AAGATACATA TAACATGAAT GTTGCCAGAA GTTCAGGGAA    1380

CTTGAATATT CTTTTGGTTC TTGGTGGAAA ATATCCAACA GATGAACAAT TTGACATTAT    1440

TTCACACTTT GATTCTAGCA ACTCTGTAAC ACCATCATGG GTTATTGTTG ATGTACATAA    1500

ATATATATTA CAAATCTGTA TACCATTGGT TCAAATTGTT ACAACATTTG TTTGAAGCAC    1560

ACCTGCAGCA ATAATACACA GGATGCAAAA CGAAGAGCGA AACTATATGA CGCCAACGAT    1620

AGACATAAAC AGTTACAGTC ATCATGAAAA CAGAATTATA TGGTACAGCA AAAATTACAC    1680

TAAGAGGCAA GAGTCTCACC GACGACGATG AGAGAGTTTA CGGTTAGACC TCTTTCCACC    1740

GGTTGATTTC GATGTGGAAG AAGTCGAATC TGTCAGGGAC GAATTCCTA ATTCCAAATT     1800

GTCCTCACTA AAGGCCTTCT TTAGTGTCTC TTGTATTTCC ATGTACCTTT GCTTCTTTTG    1860
```

```
TAGTCGTTTC TCAGCAGTGT CGTCTTCTCC GCAAGCCAGT TGAGTCAAGT CCTCACAGTT        1920

CATAATCTGG TCGAGCACTG CCGAACAGCG CGGGAAGAAT CGTTTCCCGA GTTCCACTGA        1980

TGATAAAAAA AACAAGGTCA GACAGCAAGT AACAAAACCA TGTTTAAAGA TCATTTAGTT        2040

TTGTTTTTTG TGATAAGGAG TCCGATGAAG TGGGTGAGAA TCCATACCGG TTTTAGAAAG        2100

CGCTTTTAGT CTACTTTGAT GCTCTTCTAG GATTCTGAAA GGTGCTATCT TTACACCCGG        2160

TGATGTTCTC TTCGTACCAG TGAGACGGTC AGGCTCGAGG CTAGTCACTA TGAACTCACA        2220

TGTTCCCTTC ATTTCGGCGA TCTCCATTGC AGCTTGTGCT TCCGTTGGAA AAAGACGTTG        2280

AGCAAGTGCA ACTAAACAGT GGACGACACA AAGAATAGTT ATCATTAGTT CACTCAGTTT        2340

CCTAATAGAG AGGACATAAA TTTAATTCAA ACATATAAGA AATAAGACTT GATAGATACC        2400

TCTATTTTCA AGATCGAGCA GCGTCATCTT CAATTCATCG GCCGCCACTG CAAAAGAGGG        2460

AGGAACATCT CTAGGAATTT GTTCTCGTTT GTCTTCTTGC TCTAGTATTT CTACACATAG        2520

TCGGCCTTTG AGAGAATGCT TGCATTGCTC CGGGATATTA TTACATTCAA CCGCCATAGT        2580

GGCTTGTTTT GCGATCATGA GTGCGGTTCT ACCTTCCAAA GTTGCTTCTG ATGCACTTGC        2640

ACCTTTTTCC AATAGAGATA GTATCAATTG TGGCTCCTTC CGCATCGCAG CAACATGAAG        2700

CACCGTATAT CCCCTCGGAT TCCTATGGTT GACATCGGCA AGATCAAGTT TTAAAAGATC        2760

TGTTGCGGTC TTCACATTGC AATATGCAAC AGCGAAATGA AGAGCACACG CATCATCTAG        2820

ATTGGTGTGA TCCTCTTTCA AAAGCAACTT GACTAACTCA ATATCATCCG AGTCAAGTGC        2880

CTTATGTACA TTCGAGACAT GTTTCTTTAC TTTAGGTACC TCCAAACCAA GCTCTTTACG        2940

TCTATCAATT ATCTCTTTAA CAAGCTCTTC CGGCAATGAC TTTTCAAGAC TAACCATATC        3000

TACATTAGAC TTGACAATAA TCTCTTTACA TCTATCCAAT AGCTTCATAC AAGCTTTACC        3060

ACATATATTA GCAAGCTTGA GTATAACCAA TGTGTCCTCT ATAACAACTT TGTCTACAAC        3120

GTCCAATAAG TGCCTCTGAA ATACAAATAC AAGTACTCAA GTAAGAACAT ATTCATGAAT        3180

GTGTAACCAT AGCTTAATGC AGATGGTGTT TTACCTGATA GAGAGTAATT AATTCAGGGA        3240

TCTTGAAGAT GAAAGCCAAA TAGAGAACCT CCAACATGAA ATCCACCGCC GGCCGGCAAG        3300

CCACGTGGCA GCAATTCTCG TCTGCGCATT CAGAAACTCC TTTAGGCGGC GGTCTCACTC        3360

TGCTGCTGTA AACATAAGCC AAAACAGTCA CAACCGAATC GAAACCGACT TCGTAATCCT        3420

TGGCAATCTC CTTAAGCTCG AGCTTCACGG CGGCGGTGTT GTTGGAGTCT TTCTCCTTCT        3480

TAGCGGCGGC TAAAGCGCTC TTGAAGAAAG AGCTTCTCGC TGACAAAACG CACCGGTGGA        3540

AAGAAACTTC CCGGCCGTCG GAGAGAACAA GCTTAGCGTC GCTGTAGAAA TCATCCGGCG        3600

AGTCAAAGAC GGATTCGAAG CTGTTGGAGA GCAATTGCAG AGCAGATACA TCAGGTCCGG        3660

TGAGTACTTG TTCGGCGGCC AGATAAACAA TAGAGGAGTC GGTGTTATCG GTAGCGACGA        3720

AACTAGTGCT GCTGATTTCA TAAGAATCGG CGAATCCATC AATGGTGGTG TCCATCAACA        3780

GGTTCCGATG AATTGAAATT CACAAATTAA AGAGATCTCT GCTAATCAAC GAAGAGACCT        3840
```

78

```
TATCAACTGG ATTTGGTTAA AGATCGAAGA TAACCATTGA CGAGCAGAGC CAAGTCAAGT    3900

CAACGAGAGT GGTGGTGAGA TATGAAGAAG CATCCTCGTC CCACGGTTTA CATTTCACCA    3960

AAACCGGTAA ATTTCCAGGA AAGGAATCTT TGTCAGAGAT CTTTTTTAAA AAGATATAAC    4020

AGGAAGCTAA ACCGGTTCGG GTTATAAATG TTAGTATTTA TACCGGAGAC ATTTTGTGTT    4080

GCTAATTTTT GTATATGAGA AGTTCAATCC GGTTCGGTAA GCCCCTGAAC CAAACTAGAT    4140

TTGGAGATGA TATAAATATA TAAAATTTAT TTTTCATCCG GTTCGTTATT TTCATATAAA    4200

TATATAAATA TTATTTTTTA AATTTAAGAA TTAGATTTAC ATGTGAAAGT TACATTTCTG    4260

TTTATTTTCT TTGAAGTAAA ATGATAAAGG GAACGTATAT TAAGTTTCAT GCTTTATTCA    4320

CATAAGTTTT GTAATGTATA TTATATTTTT CGTTTATTGA AAAAGTAATT TTCAGTGTTC    4380

AGCATGTTTA CACTATAATT AAATCAAGTC GAATATTTCC TGGAACTATT CTCCTTGTTC    4440

TATAGCAAAT GAAAACGCTC TTCACAACAA AATCATTATA GATATAGGAA TAAATTACAT    4500

TAAAAACATG AAAGTCATAA TGAATATATT TTTTTAATTA GGATTTGATT TAAAAACAAT    4560

TATTGTATAC ATATAAAAGA CTTCTTTAGT TATTTGCCTT CAACTTCTCG TTCTGAATCA    4620

TGCGATAAAT CAGCTTTTTC AATAACTACG ACGTAAAAGC AAATTCATAA CACGTCTAAA    4680

CAAATTTGGC TCATCCTTCA CTTGATTGGT GTTTTCCGGA CTCGATGTTG CTGGAAACTG    4740

AGAAGAAGAA GGAATCTGCA TAATCACCTC TTGGTTCCTC ACCGGTAGAC TCATTTTGTT    4800

GGATCGAAAA CGATCGAGAT CAGAAAATGA AAAGATAGGT TAAAGATGCC TATGAATACA    4860

ACAACGTAAG ATTATGTTGA ATAAACAGAG TACTTTATAT AGGAGTTATA ATAAGGTAAA    4920

TAAATTATTG CTTTCCGCGT TTTTTACTTT TGTATTTCTT AAATGATAAG TTAAATTAGG    4980

ATAAGATTTG TATGATTTTA AGTAAATTTA CAATAACTCT CTATAACTCA ATAGCATCAC    5040

ATATTTAATT AATTTTACTA ATTATCTTTT GAACAATTTT ATGAAATAGT TTTCTTTTAA    5100

TTAATTTTTT AAAATGATAT ATTATAAAAT TTAATTGAAT CAATCTGATA TAATTTTTTT    5160

ATCTTCTACC ATCTATTATA GTTGATAAAT ATTGTGATAA ACTTTAGATA AACACCCAAT    5220

TGCCAAATAT TTAATAAATT TTGTGTACCA TGCGTTTTTT TTGGAGAATA TATATACGTG    5280

GACAGCATAC CGTACATATA TTGTATAAAA GCTTATAAAA CATAGATACG GGTTATATTG    5340

GTAAGCTATA AATATATGTA AACAATAGTA AGATATTACG TGTTGTGTCT AAATATGTGT    5400

TGCTTTAGAT ATTATGTATA TCTAATATAT TAAAATATCT TTTATTAACT AATATATTAT    5460

TTAAGAGAGA AAATTGGGAC ACTATTTTCT ATACAGTAAC TGTTTTCAAC TATAAACAGG    5520

AACCCTTGAT ATAATAAAAT AACTAGCCAA AAAATCAGAT TAAATATTCA TAAAACAATG    5580

TTTGGTATTA TTACATAAAC CTAAGAAACA AAATTCAATA TTCCTTTTTA CCTTATAAAA    5640

AACAATTAAA CATCACTAGA TATATTTATG CCCCACAATG AGCGAGCCAA TTGAGACTTG    5700

AGACTTGAGA TCCTTGTCAA CTACGTTTGC ATTTGTCGGC CCATTTTTTT TATTTTTTTT    5760
```

79
```
P:\PATENT APPLICATION FILES\20000'S\21212-13\21212C\21212C SPECIFICATION.DOC
```

```
TTAAAGTGTC GGCCCGTTGC TTCTTCCGTT CAGATCAACC CTCTCGTAAT CAGAACAAAA    5820

CGGAAAACAA ACGAAAGAAC AATCAGATCC CTCTTTTTTT GCATAAACTA AATTCAACTT    5880

CTCTGCGTTT ATGTTGTAGA GGCAACCACG ATCACTACTA CGAAACAATA CAACGTCGTT    5940

GCTTGGAGTC CACGTAATCA AATCTACTCC AATGCTTTTA ATATCTTTCA CTTTAACCCA    6000

CGACTTTTCA AAACTGCTCT TTAAAACCCA TAACTCGTGA ACATCTTCTT GATCTTTGTT    6060

TGTCCACTGA CGAATAGCAC CTAGCTTCCC TTCGTATCTG ACTAATCCTG AGAAAACATC    6120

AGAGTTCGGA GTATGGAAGA AGGACCAAGT TTCGGTTTTG AGACAAAACC GGATCACATT    6180

GTTGTTCCGT GATATCCAAT GCAAGAACCC CGAAACTTGT ATCGGGTTGG AAAAAATTAA    6240

TCTGTCTGTT TTTGGTAGAC GCAAATTTTC TAATCTCTTC CAGGTAAACG AATCAGAATC    6300

GAAAACTTCG CACATAAAAG TTCTGTGATT CAAATGGTAG ATACCCCGAG ACATACACAT    6360

ACGCCGAGAC TGCGAAAGCC TTTGTATTTT ATACCGGAAA GGGTTCAATC CGATTACCGC    6420

TAAACCCAAT GACATATCCC AACCCTTCAC TTCTGGCTTT GGTATGACCT GATACTGTTT    6480

AGTGGTTGGT TTGAAGACTA TGTATCCACG TGATGGTTTT GTATACTTAA CACAAAGCAA    6540

TATCCCATGA CTTGCATCAC AAGCTTCGAT CTTTATCATT CCGGGTGGCA GAAAGTCGAT    6600

GGAGACTCCA TTGTTTTGTA AATCACTCCT CTCATGGACA AAACTGGTTC GAAGTTCGTG    6660

TCCTTTTACT ATGTAGTGTT GTATGAAGTA TCCCGAAATA CGATTGGTTC TAAGGAGATT    6720

AAGATTGACA AACCATGACT CGTAGCTTCT CTTGTTGCAC TCTTTATTCA GGAGCCTGAA    6780

TTTTCCGATT TTTGACGCCG GAAGATAAGA AAGAAATTCT TGGATCATGT CTTGATTTAT    6840

CACCGGAGAA CTCATGATCC TGTCGGGAAT AAAGAGATGA GCACGATCAC TGAATGAGAA    6900

ATGAAAAAAT CAGGATCGGT AGAGAACAAC TTATGATGAA TAAAGTGTTT ATATATCCTT    6960

TCTTTGTTTA AGGAAAGTAT CAAAATTTGC CTTTTTCTTC GCTAGTCCTA AAACAAACAA    7020

ATTAACCAAA AGATAAAATC TTTCATGATT AATGTTACTT GTGATACCTT AAGCCAAAAC    7080

TTTATCTTTA GACTTTTAAC CAAATCTACA GTAATTTAAT TGCTAGACTT AGGAAACAAC    7140

TTTTTTTTTT ACCCAACAAT CTTTGGATTT TAATTGTTTT TTTTTCTACT AATAGATTAA    7200

CAACTCATTA TATAATAATG TTTCTATCAT AATTGACAAT TCTTTCTTTT TAATAAACAT    7260

CCAGCTTGTA TAATAATCCA CAAGTCAATT TCACCATTTT GGCCAATTTA TTTTCTTATA    7320

AAAATTAGCA CAAAAAAGAT TATCATTGTT TAGCAGATTT AATTTCTAAT TAACTTACGT    7380

AATTTCCATT TTCCATAGAT TTATCTTTCT TTTTATTTCC TTAGTTATCT TAGTACTTTC    7440

TTAGTTTCCT TAGTAATTTT AAATTTTAAG ATAAATATATT GAAATTAAAA GAAGAAAAAA    7500

AACTCTAGTT ATACTTTTGT TAAATGTTTC ATCACACTAA CTAATAATTT TTTTTAGTTA    7560

AATTACAATA TATAAACACT GAAGAAAGTT TTTGGCCCAC ACTTTTTTGG GATCAATTAG    7620

TACTATAGTT AGGGGAAGAT TCTGATTTAA AGGATACCAA AAATGACTAG TTAGGACATG    7680

AATGAAAACT TATAATCTCA ATAACATACA TACGTGTTAC TGAACAATAG TAACATCTTA    7740
```

80

```
CGTGTTTTGT CCATATATTT GTTGCTTATA AATATATTCA TATAACAATG TTTGCATTAA        7800

GCTTTTAAGA AGCACAAAAC CATATAACAA AATTAAATAT TCCTATCCCT ACCAAAAAAA        7860

AAAATTAAAT ATTCCTACAG CCTTGTTGAT TATTTTATGC CCTACGTTGA GCCTTGTTGA        7920

CTAGTTTGCA TTTGTCGGTC CATTTCTTCT TCCGTCCAGA TCAACCCTCT CGTAATCAGA        7980

ACAAAAGGGG AAACAAACGT AAGAGGCAAA ATCCTTGTTT GTATGAACTA AGTTTAACTT        8040

CTCTGTGTTT AAGTTGTAGA GGCAAACATG ATCCCAACTA GAAAGCATTA CGACGTCGTT        8100

GCTTGGTATC CACGTAATAT GCTCTACTCC AATGCTTTCA ATATCTTTCA CTTTTTCCCA        8160

CGACTTTTCA AAACTGCTCT TTAAAACCCA TAATCTGTGA ACATCTTCTT GATTGTTGTT        8220

TATCCAGTGA CGAATAACAC CTAGCTTCCC TTCGTAGCTG ACTAACTCTG GGAATAAACC        8280

AACGTTTGGA GTATGTAAGA AAGACCAAGT TTCGGTTTTG GGACATAACC GGATCACATT        8340

GTGGTTCCAT GATCTCCAAT GCAAGAACCC TGAAGCTTGT ACCGGGTTTG AAAGAATTAG        8400

ACCGTCTGTT CTCGGTAGAC GCAAATTTTT TAATCTCTTC CACATAAACG AATCGGAATC        8460

AAAAACTTCG CACGCAAAAG TTCTGAGATT CCGAGTCATA CCAGGCGATT TCGAAAGCCT        8520

AAATATTTTA TACCGGAAAG GCTGCAATCC GGTTACCGTT AGACCTAATG ACTTATCACA        8580

ACTCCTCACT TTTGGGTTTG GTATGATCTG ATACTGTTTT GTTGTTGGTT TGCAGACTAT        8640

GTATTCCGGT ATTGGTCTTG TATCATTATA ACAAAGCAAT ATCCCATGAC GTGCATCACA        8700

AGCTTTGATC TTTACCTCTC CTTGTGGCAG AAAATCGATG GAGACTCCTT TGTTATCCAA        8760

ATCTCTCCTC TCATGGAAAA AACTGGTATC AAGTTTGTAT CCTCTTTCGT AGCGTTCTAG        8820

GAAGTATCCA GAGATATTGT TGGTTCGATG GAGATTTAGG TTGACAAACC AAGACTCGTA        8880

GCTTCTCTTG TTGCACTCTT TATTGATGAG CCTCAATTTT CCGATTTCGG ACCCCCGAAG        8940

ATAAGAAAGA ACCTCTTGGA TCGTGTCCTG ATTTATCACC GGAGAACTCA TGATCTTATT        9000

GGAAAAAGA AAGAAAGAGA TGAGCACGAT CAGTGAATGA GATATATAGA AATCAGGATT        9060

GGTAGAGAAC CGACGATGAT GAATATACAA GTGTTTATAA GTATCACAAA TTGCCTTTTT        9120

CTTCGCTAGT CCCAAAACAA GCAAATTAAC CAAAGATAAA ATCTTCATTA ATGTTTTCCT        9180

TTTTCTTCGC CAGTCCCAGA TAAAAATATA TATAAAATAT TTCATTAGGT TACTTGTAGT        9240

ACCTTGAGCC CAAAGTTTCT CTTTTGACTT TTAACCAAAT TAACAGTAAT TTAATAGCTA        9300

GACTTAGAAA ACAACATTTT GTATATATAT TCTTTGACAT CAAAATTCAA CAATCTTTGG        9360

GTTTCTATAG TGTTTTTTTT CTTATTCTAA TAGATTACCA CTCATTATAT CATATACAAA        9420

GTGTTTCCTT TTCAATCAAC ATCCATTTTC TTTAAAAATT AGCAAGTTTG TTCTTATATC        9480

ATCATTCAGC AGATTTCTTA ATTAAACTTA GTGATTTCCA TTTTGCACCT ATATGTTTCT        9540

CTTTCTTAGT TTAGTACTTT AAATTTTCAT ATATATAATT TATTAAAATT AAAAGTAAAA        9600

ACTCCAGTTT AACTTATGTT AAATGTTTCA TCACACTAAA AGAGCATTAA GTAATAAATA        9660
```

81

```
TTTTAGCTTT ATGAAAAAAA ATATCAAATC ACTGAAGACA TTTGTTGGCC TATACTCTAT      9720

TTTTTATTTG GCCAATTAGT AATAGACTAA TAGTAACTCA TATGATATCT CTCTAATTCT      9780

GGCGAAACGA ATATTCTGAT TCTAAAGATA GTAAAAATGA ATTTTGATGA AGGGAATACT      9840

ATTTCACACA CCTAGAAAGA GTAAGGTAGA AACCTTTTTT TTTTTGGTCA GATTCTTGTA      9900

TCAAGAAGTT CTCATCGAT                                                  9919
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5655 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

   (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (ix) FEATURE:
        (A) NAME/KEY: exon
        (B) LOCATION: 2787..3347
        (D) OTHER INFORMATION: /product= "1st exon of NIM1"

    (ix) FEATURE:
        (A) NAME/KEY: exon
        (B) LOCATION: 3427..4162
        (D) OTHER INFORMATION: /product= "2nd exon of NIM1"

    (ix) FEATURE:
        (A) NAME/KEY: exon
        (B) LOCATION: 4271..4474
        (D) OTHER INFORMATION: /product= "3rd exon of NIM1"

    (ix) FEATURE:
        (A) NAME/KEY: exon
        (B) LOCATION: 4586..4866
        (D) OTHER INFORMATION: /product= "4th exon of NIM1"

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: join(2787..3347, 3427..4162, 4271..4474, 4586..4866)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
TGTGATGCAA GTCATGGGAT ATTGCTTTGT GTTAAGTATA CAAAACCATC ACGTGGATAC        60

ATAGTCTTCA AACCAACCAC TAAACAGTAT CAGGTCATAC CAAAGCCAGA AGTGAAGGGT       120

TGGGATATGT CATTGGGTTT AGCGGTAATC GGATTGAACC CTTTCCGGTA TAAAATACAA       180

AGGCTTTCGC AGTCTCGGCG TATGTGTATG TCTCGGGGTA TCTACCATTT GAATCACAGA       240

ACTTTTATGT GCGAAGTTTT CGATTCTGAT TCGTTTACCT GGAAGAGATT AGAAAATTTG       300

CGTCTACCAA AAACAGACAG ATTAATTTTT TCCAACCCGA TACAAGTTTC GGGGTTCTTG       360

CATTGGATAT CACGGAACAA CAATGTGATC CGGTTTTGTC TCAAAACCGA AACTTGGTCC       420
```

<div align="center">82</div>

```
TTCTTCCATA CTCCGAACTC TGATGTTTTC TCAGGATTAG TCAGATACGA AGGGAAGCTA        480

GGTGCTATTC GTCAGTGGAC AAACAAAGAT CAAGAAGATG TTCACGAGTT ATGGGTTTTA        540

AAGAGCAGTT TTGAAAAGTC GTGGGTTAAA GTGAAAGATA TTAAAAGCAT TGGAGTAGAT        600

TTGATTACGT GGACTCCAAG CAACGACGTT GTATTGTTTC GTAGTAGTGA TCGTGGTTGC        660

CTCTACAACA TAAACGCAGA GAAGTTGAAT TTAGTTTATG CAAAAAAAGA GGGATCTGAT        720

TGTTCTTTCG TTTGTTTTCC GTTTTGTTCT GATTACGAGA GGGTTGATCT GAACGGAAGA        780

AGCAACGGGC CGACACTTTA AAAAAAAAAT AAAAAAAATG GCCGACAAA TGCAAACGTA         840

GTTGACAAGG ATCTCAAGTC TCAAGTCTCA ATTGGCTCGC TCATTGTGGG GCATAAATAT        900

ATCTAGTGAT GTTTAATTGT TTTTTATAAG GTAAAAAGGA ATATTGAATT TTGTTTCTTA        960

GGTTTATGTA ATAATACCAA ACATTGTTTT ATGAATATTT AATCTGATTT TTTGGCTAGT       1020

TATTTTATTA TATCAAGGGT TCCTGTTTAT AGTTGAAAAC AGTTACTGTA TAGAAAATAG       1080

TGTCCCAATT TTCTCTCTTA AATAATATAT TAGTTAATAA AAGATATTTT AATATATTAG       1140

ATATACATAA TATCTAAAGC AACACATATT TAGACACAAC ACGTAATATC TTACTATTGT       1200

TTACATATAT TTATAGCTTA CCAATATAAC CCGTATCTAT GTTTTATAAG CTTTTATACA       1260

ATATATGTAC GGTATGCTGT CCACGTATAT ATATTCTCCA AAAAAAACGC ATGGTACACA       1320

AAATTTATTA AATATTTGGC AATTGGGTGT TTATCTAAAG TTTATCACAA TATTTATCAA       1380

CTATAATAGA TGGTAGAAGA TAAAAAAATT ATATCAGATT GATTCAATTA AATTTTATAA       1440

TATATCATTT TAAAAAATTA ATTAAAAGAA AACTATTTCA TAAAATTGTT CAAAAGATAA       1500

TTAGTAAAAT TAATTAAATA TGTGATGCTA TTGAGTTATA GAGAGTTATT GTAAATTTAC       1560

TTAAAATCAT ACAAATCTTA TCCTAATTTA ACTTATCATT TAAGAAATAC AAAAGTAAAA       1620

AACGCGGAAA GCAATAATTT ATTTACCTTA TTATAACTCC TATATAAAGT ACTCTGTTTA       1680

TTCAACATAA TCTTACGTTG TTGTATTCAT AGGCATCTTT AACCTATCTT TTCATTTTCT       1740

GATCTCGATC GTTTTCGATC CAACAAAATG AGTCTACCGG TGAGGAACCA AGAGGTGATT       1800

ATGCAGATTC CTTCTTCTTC TCAGTTTCCA GCAACATCGA GTCCGGAAAA CACCAATCAA       1860

GTGAAGGATG AGCCAAATTT GTTTAGACGT GTTATGAATT TGCTTTTACG TCGTAGTTAT       1920

TGAAAAAGCT GATTTATCGC ATGATTCAGA ACGAGAAGTT GAAGGCAAAT AACTAAAGAA       1980

GTCTTTTATA TGTATACAAT AATTGTTTTT AAATCAAATC CTAATTAAAA AAATATATTC       2040

ATTATGACTT TCATGTTTTT AATGTAATTT ATTCCTATAT CTATAATGAT TTTGTTGTGA       2100

AGAGCGTTTT CATTTGCTAT AGAACAAGGA GAATAGTTCC AGGAAATATT CGACTTGATT       2160

TAATTATAGT GTAAACATGC TGAACACTGA AAATTACTTT TTCAATAAAC GAAAAATATA       2220

ATATACATTA CAAAACTTAT GTGAATAAAG CATGAAACTT AATATACGTT CCCTTTATCA       2280

TTTTACTTCA AAGAAAATAA ACAGAAATGT AACTTTCACA TGTAAATCTA ATTCTTAAAT       2340
```

83

```
TTAAAAAATA ATATTTATAT ATTTATATGA AAATAACGAA CCGGATGAAA AATAAATTTT      2400

ATATATTTAT ATCATCTCCA AATCTAGTTT GGTTCAGGGG CTTACCGAAC CGGATTGAAC      2460

TTCTCATATA CAAAAATTAG CAACACAAAA TGTCTCCGGT ATAAATACTA ACATTTATAA      2520

CCCGAACCGG TTTAGCTTCC TGTTATATCT TTTTAAAAAA GATCTCTGAC AAAGATTCCT      2580

TTCCTGGAAA TTTACCGGTT TTGGTGAAAT GTAAACCGTG GGACGAGGAT GCTTCTTCAT      2640

ATCTCACCAC CACTCTCGTT GACTTGACTT GGCTCTGCTC GTCAATGGTT ATCTTCGATC      2700

TTTAACCAAA TCCAGTTGAT AAGGTCTCTT CGTTGATTAG CAGAGATCTC TTTAATTTGT      2760

GAATTTCAAT TCATCGGAAC CTGTTG ATG GAC ACC ACC ATT GAT GGA TTC GCC      2813
                             Met Asp Thr Thr Ile Asp Gly Phe Ala
                              1               5

GAT TCT TAT GAA ATC AGC AGC ACT AGT TTC GTC GCT ACC GAT AAC ACC      2861
Asp Ser Tyr Glu Ile Ser Ser Thr Ser Phe Val Ala Thr Asp Asn Thr
 10              15              20              25

GAC TCC TCT ATT GTT TAT CTG GCC GCC GAA CAA GTA CTC ACC GGA CCT      2909
Asp Ser Ser Ile Val Tyr Leu Ala Ala Glu Gln Val Leu Thr Gly Pro
            30              35              40

GAT GTA TCT GCT CTG CAA TTG CTC TCC AAC AGC TTC GAA TCC GTC TTT      2957
Asp Val Ser Ala Leu Gln Leu Leu Ser Asn Ser Phe Glu Ser Val Phe
            45              50              55

GAC TCG CCG GAT GAT TTC TAC AGC GAC GCT AAG CTT GTT CTC TCC GAC      3005
Asp Ser Pro Asp Asp Phe Tyr Ser Asp Ala Lys Leu Val Leu Ser Asp
            60              65              70

GGC CGG GAA GTT TCT TTC CAC CGG TGC GTT TTG TCA GCG AGA AGC TCT      3053
Gly Arg Glu Val Ser Phe His Arg Cys Val Leu Ser Ala Arg Ser Ser
 75              80              85

TTC TTC AAG AGC GCT TTA GCC GCC GCT AAG AAG GAG AAA GAC TCC AAC      3101
Phe Phe Lys Ser Ala Leu Ala Ala Ala Lys Lys Glu Lys Asp Ser Asn
 90              95              100             105

AAC ACC GCC GCC GTG AAG CTC GAG CTT AAG GAG ATT GCC AAG GAT TAC      3149
Asn Thr Ala Ala Val Lys Leu Glu Leu Lys Glu Ile Ala Lys Asp Tyr
                110             115             120

GAA GTC GGT TTC GAT TCG GTT GTG ACT GTT TTG GCT TAT GTT TAC AGC      3197
Glu Val Gly Phe Asp Ser Val Val Thr Val Leu Ala Tyr Val Tyr Ser
            125             130             135

AGC AGA GTG AGA CCG CCG CCT AAA GGA GTT TCT GAA TGC GCA GAC GAG      3245
Ser Arg Val Arg Pro Pro Pro Lys Gly Val Ser Glu Cys Ala Asp Glu
            140             145             150

AAT TGC TGC CAC GTG GCT TGC CGG CCG GCG GTG GAT TTC ATG TTG GAG      3293
Asn Cys Cys His Val Ala Cys Arg Pro Ala Val Asp Phe Met Leu Glu
 155             160             165

GTT CTC TAT TTG GCT TTC ATC TTC AAG ATC CCT GAA TTA ATT ACT CTC      3341
Val Leu Tyr Leu Ala Phe Ile Phe Lys Ile Pro Glu Leu Ile Thr Leu
170             175             180             185

TAT CAG GTAAACACC ATCTGCATTA AGCTATGGTT ACACATTCAT GAATATGTTC      3397
Tyr Gln
```

84

```
TTACTTGAGT ACTTGTATTT GTATTTCAG AGG CAC TTA TTG GAC GTT GTA GAC        3450
                                   Arg His Leu Leu Asp Val Val Asp
                                       190                     195

AAA GTT GTT ATA GAG GAC ACA TTG GTT ATA CTC AAG CTT GCT AAT ATA        3498
Lys Val Val Ile Glu Asp Thr Leu Val Ile Leu Lys Leu Ala Asn Ile
                200                 205                 210

TGT GGT AAA GCT TGT ATG AAG CTA TTG GAT AGA TGT AAA GAG ATT ATT        3546
Cys Gly Lys Ala Cys Met Lys Leu Leu Asp Arg Cys Lys Glu Ile Ile
            215                 220                 225

GTC AAG TCT AAT GTA GAT ATG GTT AGT CTT GAA AAG TCA TTG CCG GAA        3594
Val Lys Ser Asn Val Asp Met Val Ser Leu Glu Lys Ser Leu Pro Glu
        230                 235                 240

GAG CTT GTT AAA GAG ATA ATT GAT AGA CGT AAA GAG CTT GGT TTG GAG        3642
Glu Leu Val Lys Glu Ile Ile Asp Arg Arg Lys Glu Leu Gly Leu Glu
        245                 250                 255

GTA CCT AAA GTA AAG AAA CAT GTC TCG AAT GTA CAT AAG GCA CTT GAC        3690
Val Pro Lys Val Lys Lys His Val Ser Asn Val His Lys Ala Leu Asp
260                 265                 270                 275

TCG GAT GAT ATT GAG TTA GTC AAG TTG CTT TTG AAA GAG GAT CAC ACC        3738
Ser Asp Asp Ile Glu Leu Val Lys Leu Leu Leu Lys Glu Asp His Thr
                280                 285                 290

AAT CTA GAT GAT GCG TGT GCT CTT CAT TTC GCT GTT GCA TAT TGC AAT        3786
Asn Leu Asp Asp Ala Cys Ala Leu His Phe Ala Val Ala Tyr Cys Asn
            295                 300                 305

GTG AAG ACC GCA ACA GAT CTT TTA AAA CTT GAT CTT GCC GAT GTC AAC        3834
Val Lys Thr Ala Thr Asp Leu Leu Lys Leu Asp Leu Ala Asp Val Asn
            310                 315                 320

CAT AGG AAT CCG AGG GGA TAT ACG GTG CTT CAT GTT GCT GCG ATG CGG        3882
His Arg Asn Pro Arg Gly Tyr Thr Val Leu His Val Ala Ala Met Arg
        325                 330                 335

AAG GAG CCA CAA TTG ATA CTA TCT CTA TTG GAA AAA GGT GCA AGT GCA        3930
Lys Glu Pro Gln Leu Ile Leu Ser Leu Leu Glu Lys Gly Ala Ser Ala
340                 345                 350                 355

TCA GAA GCA ACT TTG GAA GGT AGA ACC GCA CTC ATG ATC GCA AAA CAA        3978
Ser Glu Ala Thr Leu Glu Gly Arg Thr Ala Leu Met Ile Ala Lys Gln
                360                 365                 370

GCC ACT ATG GCG GTT GAA TGT AAT AAT ATC CCG GAG CAA TGC AAG CAT        4026
Ala Thr Met Ala Val Glu Cys Asn Asn Ile Pro Glu Gln Cys Lys His
                375                 380                 385

TCT CTC AAA GGC CGA CTA TGT GTA GAA ATA CTA GAG CAA GAA GAC AAA        4074
Ser Leu Lys Gly Arg Leu Cys Val Glu Ile Leu Glu Gln Glu Asp Lys
            390                 395                 400

CGA GAA CAA ATT CCT AGA GAT GTT CCT CCC TCT TTT GCA GTG GCG GCC        4122
Arg Glu Gln Ile Pro Arg Asp Val Pro Pro Ser Phe Ala Val Ala Ala
        405                 410                 415

GAT GAA TTG AAG ATG ACG CTG CTC GAT CTT GAA AAT AGA G                  4162
Asp Glu Leu Lys Met Thr Leu Leu Asp Leu Glu Asn Arg
420                 425                 430
```

GTATCTATCA AGTCTTATTT CTTATATGTT TGAATTAAAT TTATGTCCTC TCTATTAGGA   4222

AACTGAGTGA ACTAATGATA ACTATTCTTT GTGTCGTCCA CTGTTTAG   TT GCA CTT   4278
                                                          Val Ala Leu
                                                              435

GCT CAA CGT CTT TTT CCA ACG GAA GCA CAA GCT GCA ATG GAG ATC GCC   4326
Ala Gln Arg Leu Phe Pro Thr Glu Ala Gln Ala Ala Met Glu Ile Ala
                440   445   450

GAA ATG AAG GGA ACA TGT GAG TTC ATA GTG ACT AGC CTC GAG CCT GAC   4374
Glu Met Lys Gly Thr Cys Glu Phe Ile Val Thr Ser Leu Glu Pro Asp
                455   460   465

CGT CTC ACT GGT ACG AAG AGA ACA TCA CCG GGT GTA AAG ATA GCA CCT   4422
Arg Leu Thr Gly Thr Lys Arg Thr Ser Pro Gly Val Lys Ile Ala Pro
            470   475   480

TTC AGA ATC CTA GAA GAG CAT CAA AGT AGA CTA AAA GCG CTT TCT AAA   4470
Phe Arg Ile Leu Glu Glu His Gln Ser Arg Leu Lys Ala Leu Ser Lys
    485   490   495

ACC G GTATGGATTC TCACCCACTT CATCGGACTC CTTATCACAA AAAACAAAAC   4524
Thr
500

TAAATGATCT TTAAACATGG TTTTGTTACT TGCTGTCTGA CCTTGTTTTT TTTATCATCA   4584

G TG GAA CTC GGG AAA CGA TTC TTC CCG CGC TGT TCG GCA GTG CTC   4629
   Val Glu Leu Gly Lys Arg Phe Phe Pro Arg Cys Ser Ala Val Leu
               505   510   515

GAC CAG ATT ATG AAC TGT GAG GAC TTG ACT CAA CTG GCT TGC GGA GAA   4677
Asp Gln Ile Met Asn Cys Glu Asp Leu Thr Gln Leu Ala Cys Gly Glu
                520   525   530

GAC GAC ACT GCT GAG AAA CGA CTA CAA AAG AAG CAA AGG TAC ATG GAA   4725
Asp Asp Thr Ala Glu Lys Arg Leu Gln Lys Lys Gln Arg Tyr Met Glu
            535   540   545

ATA CAA GAG ACA CTA AAG AAG GCC TTT AGT GAG GAC AAT TTG GAA TTA   4773
Ile Gln Glu Thr Leu Lys Lys Ala Phe Ser Glu Asp Asn Leu Glu Leu
        550   555   560

GGA AAT TCG TCC CTG ACA GAT TCG ACT TCT TCC ACA TCG AAA TCA ACC   4821
Gly Asn Ser Ser Leu Thr Asp Ser Thr Ser Ser Thr Ser Lys Ser Thr
    565   570   575

GGT GGA AAG AGG TCT AAC CGT AAA CTC TCT CAT CGT CGT CGG TGA   4866
Gly Gly Lys Arg Ser Asn Arg Lys Leu Ser His Arg Arg Arg *
580   585   590

GACTCTTGCC TCTTAGTGTA ATTTTTGCTG TACCATATAA TTCTGTTTTC ATGATGACTG   4926

TAACTGTTTA TGTCTATCGT TGGCGTCATA TAGTTTCGCT CTTCGTTTTG CATCCTGTGT   4986

ATTATTGCTG CAGGTGTGCT TCAAACAAAT GTTGTAACAA TTTGAACCAA TGGTATACAG   5046

ATTTGTAATA TATATTTATG TACATCAACA ATAACCCATG ATGGTGTTAC AGAGTTGCTA   5106

GAATCAAAGT GTGAAATAAT GTCAAATTGT TCATCTGTTG GATATTTTCC ACCAAGAACC   5166

AAAAGAATAT TCAAGTTCCC TGAACTTCTG GCAACATTCA TGTTATATGT ATCTTCCTAA   5226

TTCTTCCTTT AACCTTTTGT AACTCGAATT ACACAGCAAG TTAGTTTCAG GTCTAGAGAT   5286

```
AAGAGAACAC TGAGTGGGCG TGTAAGGTGC ATTCTCCTAG TCAGCTCCAT TGCATCCAAC      5346

ATTTGTGAAT GACACAAGTT AACAATCCTT TGCACCATTT CTGGGTGCAT ACATGGAAAC      5406

TTCTTCGATT GAAACTTCCC ACATGTGCAG GTGCGTTCGC TGTCACTGAT AGACCAAGAG      5466

ACTGAAAGCT TTCACAAATT GCCCTCAAAT CTTCTGTTTC TATCGTCATG ACTCCATATC      5526

TCCGACCACT GGTCATGAGC CAGAGCCCAC TGATTTTGAG GGAATTGGGC TAACCATTTC      5586

CGAGCTTCTG AGTCCTTCTT TTTGATGTCC TTTATGTAGG AATCAAATTC TTCCTTCTGA      5646

CTTGTGGAT                                                              5655
```

(2) INFORMATION FOR SEQ ID NO:3:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 594 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Asp Thr Thr Ile Asp Gly Phe Ala Asp Ser Tyr Glu Ile Ser Ser
1               5                   10                  15

Thr Ser Phe Val Ala Thr Asp Asn Thr Asp Ser Ser Ile Val Tyr Leu
            20                  25                  30

Ala Ala Glu Gln Val Leu Thr Gly Pro Asp Val Ser Ala Leu Gln Leu
        35                  40                  45

Leu Ser Asn Ser Phe Glu Ser Val Phe Asp Ser Pro Asp Asp Phe Tyr
    50                  55                  60

Ser Asp Ala Lys Leu Val Leu Ser Asp Gly Arg Glu Val Ser Phe His
65              70                  75                  80

Arg Cys Val Leu Ser Ala Arg Ser Ser Phe Phe Lys Ser Ala Leu Ala
                85                  90                  95

Ala Ala Lys Lys Glu Lys Asp Ser Asn Asn Thr Ala Ala Val Lys Leu
            100                 105                 110

Glu Leu Lys Glu Ile Ala Lys Asp Tyr Glu Val Gly Phe Asp Ser Val
        115                 120                 125

Val Thr Val Leu Ala Tyr Val Tyr Ser Ser Arg Val Arg Pro Pro Pro
    130                 135                 140

Lys Gly Val Ser Glu Cys Ala Asp Glu Asn Cys Cys His Val Ala Cys
145             150                 155                 160

Arg Pro Ala Val Asp Phe Met Leu Glu Val Leu Tyr Leu Ala Phe Ile
                165             170                 175

Phe Lys Ile Pro Glu Leu Ile Thr Leu Tyr Gln Arg His Leu Leu Asp
            180                 185                 190

Val Val Asp Lys Val Val Ile Glu Asp Thr Leu Val Ile Leu Lys Leu
    195                 200                 205
```

87

```
Ala Asn Ile Cys Gly Lys Ala Cys Met Lys Leu Leu Asp Arg Cys Lys
    210             215             220

Glu Ile Ile Val Lys Ser Asn Val Asp Met Val Ser Leu Glu Lys Ser
225             230             235             240

Leu Pro Glu Glu Leu Val Lys Glu Ile Ile Asp Arg Arg Lys Glu Leu
            245             250             255

Gly Leu Glu Val Pro Lys Val Lys Lys His Val Ser Asn Val His Lys
            260             265             270

Ala Leu Asp Ser Asp Asp Ile Glu Leu Val Lys Leu Leu Leu Lys Glu
            275             280             285

Asp His Thr Asn Leu Asp Asp Ala Cys Ala Leu His Phe Ala Val Ala
    290             295             300

Tyr Cys Asn Val Lys Thr Ala Thr Asp Leu Leu Lys Leu Asp Leu Ala
305             310             315             320

Asp Val Asn His Arg Asn Pro Arg Gly Tyr Thr Val Leu His Val Ala
            325             330             335

Ala Met Arg Lys Glu Pro Gln Leu Ile Leu Ser Leu Leu Glu Lys Gly
            340             345             350

Ala Ser Ala Ser Glu Ala Thr Leu Glu Gly Arg Thr Ala Leu Met Ile
            355             360             365

Ala Lys Gln Ala Thr Met Ala Val Glu Cys Asn Asn Ile Pro Glu Gln
    370             375             380

Cys Lys His Ser Leu Lys Gly Arg Leu Cys Val Glu Ile Leu Glu Gln
385             390             395             400

Glu Asp Lys Arg Glu Gln Ile Pro Arg Asp Val Pro Pro Ser Phe Ala
            405             410             415

Val Ala Ala Asp Glu Leu Lys Met Thr Leu Leu Asp Leu Glu Asn Arg
            420             425             430

Val Ala Leu Ala Gln Arg Leu Phe Pro Thr Glu Ala Gln Ala Ala Met
            435             440             445

Glu Ile Ala Glu Met Lys Gly Thr Cys Glu Phe Ile Val Thr Ser Leu
    450             455             460

Glu Pro Asp Arg Leu Thr Gly Thr Lys Arg Thr Ser Pro Gly Val Lys
465             470             475             480

Ile Ala Pro Phe Arg Ile Leu Glu Glu His Gln Ser Arg Leu Lys Ala
            485             490             495

Leu Ser Lys Thr Val Glu Leu Gly Lys Arg Phe Phe Pro Arg Cys Ser
            500             505             510

Ala Val Leu Asp Gln Ile Met Asn Cys Glu Asp Leu Thr Gln Leu Ala
    515             520             525

Cys Gly Glu Asp Asp Thr Ala Glu Lys Arg Leu Gln Lys Lys Gln Arg
    530             535             540

Tyr Met Glu Ile Gln Glu Thr Leu Lys Lys Ala Phe Ser Glu Asp Asn
```

88

```
Leu Glu Leu Gly Asn Ser Ser Leu Thr Asp Ser Thr Ser Ser Thr Ser
                565                 570                 575

Lys Ser Thr Gly Gly Lys Arg Ser Asn Arg Lys Leu Ser His Arg Arg
        580                 585                 590

Arg *
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 41 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Ile Arg Arg Met Arg Arg Ala Leu Asp Ala Ala Asp Ile Glu Leu Val
1               5               10                  15

Lys Leu Met Val Met Gly Glu Gly Leu Asp Leu Asp Asp Ala Leu Ala
            20              25                  30

Val His Tyr Ala Val Gln His Cys Asn
        35              40
```

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 38 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Pro Thr Gly Lys Thr Ala Leu His Leu Ala Ala Glu Met Val Ser Pro
1               5               10                  15

Asp Met Val Ser Val Leu Leu Asp His His Ala Asp Xaa Asn Phe Arg
            20                  25                  30

Thr Xaa Asp Gly Val Thr
        35
```

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 41 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ile Arg Arg Met Arg Arg Ala Leu Asp Ala Ala Asp Ile Glu Leu Val
1   5   10   15

Lys Leu Met Val Met Gly Glu Gly Leu Asp Leu Asp Asp Ala Leu Ala
        20   25   30

Val His Tyr Ala Val Gln His Cys Asn
        35   40

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Arg Arg Pro Asp Ser Lys Thr Ala Leu His Leu Ala Ala Glu Met Val
1   5   10   15

Ser Pro Asp Met Val Ser Val Leu Leu Asp Gln
        20   25

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 41 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: not relevant
        (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ile Arg Arg Met Arg Arg Ala Leu Asp Ala Ala Asp Ile Glu Leu Val
1   5   10   15

Lys Leu Met Val Met Gly Glu Gly Leu Asp Leu Asp Asp Ala Leu Ala
        20   25   30

Val His Tyr Ala Val Gln His Cys Asn
        35   40

(2) INFORMATION FOR SEQ ID NO:9:

90

(i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 27 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: not relevant
            (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

    Arg Arg Pro Asp Ser Lys Thr Ala Leu His Leu Ala Ala Glu Met Val
    1   5   10   15

    Ser Pro Asp Met Val Ser Val Leu Leu Asp Gln
            20   25

(2) INFORMATION FOR SEQ ID NO:10:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 41 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: not relevant
            (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

    Ile Arg Arg Met Arg Arg Ala Leu Asp Ala Ala Asp Ile Glu Leu Val
    1   5   10   15

    Lys Leu Met Val Met Gly Glu Gly Leu Asp Leu Asp Asp Ala Leu Ala
            20   25   30

    Val His Tyr Ala Val Gln His Cys Asn
            35   40

(2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 19 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: not relevant
            (D) TOPOLOGY: not relevant

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

    Pro Thr Gly Lys Thr Ala Leu His Leu Ala Ala Glu Met Val Ser Pro
    1   5   10   15

    Asp Met Val

(2) INFORMATION FOR SEQ ID NO:12:

91

(i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "oligonucleotide"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AATTCTAAAG CATGCCGATC GG   22

(2) INFORMATION FOR SEQ ID NO:13:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 21 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "oligonucleotide"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTCCGATC GGCATGCTTT A   21

(2) INFORMATION FOR SEQ ID NO:14:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 22 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "oligonucleotide"

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AATTCTAAAC CATGGCGATC GG   22

(2) INFORMATION FOR SEQ ID NO:15:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 21 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: other nucleic acid
            (A) DESCRIPTION: /desc = "oligonucleotide"

92

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:415:

AATTCCGATC GCCATGGTTT A   21

(2) INFORMATION FOR SEQ ID NO:16:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 15 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
             (A) DESCRIPTION: /desc = "oligonucleotide"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCAGCTGGAA TTCCG   15

(2) INFORMATION FOR SEQ ID NO:17:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 19 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: other nucleic acid
             (A) DESCRIPTION: /desc = "oligonucleotide"

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGAATTCCA GCTGGCATG   19

93